Systematic Development of Sandwich Immunoassays for the Plasma Secretome

Häussler, Ragna S.; Bendes, Annika; Iglesias, Maria Jesus; Sanchez-Rivera, Laura; Dodig-Crnkovi, Tea; Byström, Sanna; Fredolini, Claudia; Birgersson, Elin; Dale, Matilda; Edfors, Fredrik

Total number of authors:
16

Published in:
Proteomics

Link to article, DOI:
10.1002/pmic.201900008

Publication date:
2019

Document Version
Publisher's PDF, also known as Version of record

Systematic Development of Sandwich Immunoassays for the Plasma Secretome

*Ragna S. Häussler, Annika Bendes, MariaJesus Iglesias, Laura Sanchez-Rivera, Tea Dodig-Crnković, Sanna Byström, Claudia Fredolini, Elin Birgersson, Matilda Dale, Fredrik Edfors, Linn Fagerberg, Johan Rockberg, Hanna Tegel, Mathias Uhlén, Ulrika Qundos, and Jochen M. Schwenk**

The plasma proteome offers a clinically useful window into human health. Recent advances from highly multiplexed assays now call for appropriate pipelines to validate individual candidates. Here, a workflow is developed to build dual binder sandwich immunoassays (SIA) and for proteins predicted to be secreted into plasma. Utilizing suspension bead arrays, ≈1800 unique antibody pairs are first screened against 209 proteins with recombinant proteins as well as EDTA plasma. Employing 624 unique antibodies, dilution-dependent curves in plasma and concentration-dependent curves of full-length proteins for 102 (49%) of the targets are obtained. For 22 protein assays, the longitudinal, interindividual, and technical performance is determined in a set of plasma samples collected from 18 healthy subjects every third month over 1 year. Finally, 14 of these assays are compared with with SIAs composed of other binders, proximity extension assays, and affinity-free targeted mass spectrometry. The workflow provides a multiplexed approach to screen for SIA pairs that suggests using at least three antibodies per target. This design is applicable for a wider range of targets of the plasma proteome, and the assays can be applied for discovery but also to validate emerging candidates derived from other platforms.

1. Introduction

There is a continuously great interest in increasing our understanding about those proteins that are expected to be present in blood and found outside the intracellular space, and to apply appropriate tools to discover and validate these in a given study context.[1] Such efforts preferably target the proteins that are actively secreted in comparison to those that appear in blood due to leakage, cell death, or cellular turnaround. Today, the human secretome can be defined by bioinformatics tools annotating our genome based on sequences found in the protein-encoding regions.[2] Using an updated annotation,[3] more than 2600 proteins were defined as the secretome. Of these only about 600 proteins are predicted to be actively secreted to the

Dr. R. S. Häussler, A. Bendes, T. Dodig-Crnković, Dr. S. Byström, Dr. C. Fredolini, E. Birgersson, M. Dale, Prof. J. M. Schwenk
Division of Affinity Proteomics
Science for Life Laboratory
KTH - Royal Institute of Technology
Box 1031, 171 21 Solna, Sweden
E-mail: jochen.schwenk@scilifelab.se

Dr. M. Iglesias, Dr. L. Sanchez-Rivera
Division of Cellular and Clinical Proteomics
Science for Life Laboratory
KTH - Royal Institute of Technology
Box 1031, 171 21 Solna, Sweden

Dr. M. Iglesias
K.G. Jebsen – Thrombosis Research and Expertise Center (TREC)
Department of Clinical Medicine
UiT - The Arctic University of Norway
9010 Tromsø, Norway

The ORCID identification number(s) for the author(s) of this article can be found under https://doi.org/10.1002/pmic.201900008

DOI: 10.1002/pmic.201900008

Dr. M. Iglesias
Division of Internal Medicine
University Hospital of North Norway
9010 Tromsø, Norway

Dr. F. Edfors, Dr. L. Fagerberg, Prof. M. Uhlén
Division of Systems Biology
Science for Life Laboratory
KTH - Royal Institute of Technology
Box 1031, 171 21 Solna, Sweden

Prof. J. Rockberg, Dr. H. Tegel
Division of Protein Technology
Department of Protein Science
KTH - Royal Institute of Technology
106 91 Stockholm, Sweden

Prof. M. Uhlén
Novo Nordisk Foundation Center for Biosustainability
Technical University of Denmark
2970 Hørsholm, Denmark

Dr. U. Qundos
Atlas Antibodies AB
168 69 Bromma, Sweden

blood while another 1000 proteins were localized to the membrane and/or the intracellular space.[3]

The technically measurable content of the human plasma proteome currently contains nearly 5000 proteins when combining the efforts conducted with mass spectrometry (MS) as well as immunoassay platforms.[4] MS alone detected about 3500 plasma proteins of which a third (≈ 1230) were annotated as being secreted. It has been shown that untargeted and MS-based approaches contribute to this list with primarily cellular components. Immunoassays on the other hand are often more sensitive to detect low abundant proteins, defined as those below 1 ng mL^{-1}, and these are often linked to cytokines and inflammation processes.[5] However, and upon excluding the recent large-scale aptamer studies, only about one-third of the currently annotated 2600 proteins of the secretome[3] can be measured in plasma using other immunoassay-based methods. Plasma profiling efforts using shotgun MS, such as those by Mann and co-workers, detected 1200 proteins[6] in plasma. The latest versions of multiplexed immunoassay, not included in the above stated collection of plasma proteins, used 5000 aptamers to profile 5000 donors, as shown by Emilsson et al.[7]

Here, we present a complementary approach to multiplexed assay systems and systematically build SIAs for the proteins of the plasma secretome. Our efforts are centered around the feasibility of screening and validating the antibody (Ab) pairs for a variety of proteins at the same time, rather than focusing on only a few shortlisted targets. This pipeline is built in a way to include all secreted proteins, and hence offers an opportunity to develop assays to detect those ≈1370 secreted proteins that cannot be detected by MS today, for example, IL6. Hence it expands on previous workflows that primarily work on selected candidates.[8] Our approach was accelerated by accessing a large resource of Abs from the Human Protein Atlas (HPA) and full-length proteins generated within the Human Secretome Project[3] (HSP) within the Wallenberg Centre for Protein Research (WCPR). The study was conducted on a multiplexed bead array platform and combined the assessment of Ab pairs using both recombinant proteins and EDTA plasma. We did not preselect the secreted proteins based on other prior knowledge or particular interest but rather availability of reagents to conduct this proof-of-concept study from screening, via validation to the analysis of longitudinal samples.

Significance Statement

There is a substantial interest in monitoring human health via the analysis of accessible sample material such as blood plasma. In addition to efforts focusing on measuring more proteins in a larger number of samples, we present a complementary pipeline that allows validating the novel candidates in a systematic and quantitative manner. Since sandwich-based assays are one of the most important test concepts for clinical care, we developed our workflow to build SIAs for several secreted proteins in parallel. We also present the challenges and solutions when developing these dual binder assays and we put forward decision criteria for prioritizing antibody pairs. Furthermore, we evaluate the performance of the SIA assays in longitudinal samples, supporting the notion that protein in plasma are stable over time when being analyzed in consecutively collected samples. Lastly, we also provide a direct comparison of our data with other types of proteomics measurements, hence give valuable insights when translating assays from one technology to another, provide matching antibody pairs to develop of assays toward clinical utility or for the adaptation to other immunoassay-based technologies.

2. Experimental Section

2.1. Plasma Samples

All methods were conducted according to the Declaration of Helsinki, which establishes the regulations and guidelines for research project execution for human health.

The screening for SIA pairs was conducted on pools of anonymous donors and did not require sensitive personal information about the donors. The research did not include any type of intervention, surgery, or treatment. The Ethical Review Board in Uppsala (Dnr 2009/019) deemed that this research was not subjected to formal ethical review and approval. Samples of human K2 EDTA plasma were purchased on two occasions from Sera Laboratories International Ltd (HMPLEDTA2, now part of BioIVT, West Sussex, UK), who collects samples under IRB-approved protocols in use at their FDA-licensed donor centers with written informed consent obtained from all donors. The pools of plasma samples were generated by the supplier from mixing plasma from donors of which 50% were females.

The selected SIA pairs were then used to study samples collected from 18 subjects over a 1-year time period. Each subject consenting to participate in the longitudinal Swedish SCAPIS SciLifeLab Wellness Profiling (S3WP) program (ethical approval Dnr 407–15) donated plasma every third month. Within this study, a total of 101 healthy subjects were recruited from the ongoing Swedish CArdioPulmonary bioImage Study (SCAPIS), which is a prospective observational study of randomly selected subjects aged 50–64 years from the general Swedish population. All participants have been extensively phenotyped before entering the S3WP program.[9] Blood plasma samples were collected using 6 mL EDTA tubes (Vacuette, 456243) and centrifuged at 3000 rpm at room temperature immediately after sample collection. Then, the plasma was transferred to 0.5 mL tubes (Sarstedt, 72.730.003) and was frozen within 20 min past centrifugation. From the 101 participants, we analyzed a gender matched subset of plasma samples collected from nine male (mean age 57.6) and nine female donors (mean age 57.3). The SIA pair targeting EFEMP1 was conducted on a different selection of 18 subjects, due to the available sample volume. For this sample selection again nine male (mean age 56.6) and nine female donors (mean age 59) were chosen.

2.2. Target Selection and Generation

Protein targets for the secretome were selected according to availability of full-length proteins within the HSP and Abs from the HPA as well as considering the recombinant

protein concentration. HPA Abs needed to have a concentration of > 0.05 mg mL^{-1} for being chosen as capture antibody (capAb) and of > 0.1 mg mL^{-1} for being considered as detection Ab (detAb).

2.3. Production and Purification of Secreted Proteins

Secreted proteins were defined based on data in the Uniprot database as well as signal peptide and transmembrane region predictions made for the transcripts in the Ensembl database. A generic expression cassette, based on the CMV promoter and with an N-terminal CD33 signal peptide for secretion of all produced proteins and a C-terminal human protein C tag for purification, was used. All secreted proteins were produced by using the transient Icosagen Cell Factory system with CHOEBNALT-85 cells and the QMCF Technology (Icosagen Cell Factory OÜ, Tartu, Estonia). Cells were maintained in a 50:50 mixture of 293 SFM II (Gibco, 11686029) and CD CHO medium (Gibco, 10743001) with a supplement of 6 mM GlutaMAX (Gibco, 35050061) and 10 mL per l HT supplement 50X (Gibco, 41065012) at 37 °C on an orbital shaker. A total of six million cells were transfected by electroporation. The transfected cells were added to fresh prewarmed 20 mL medium containing penicillin–streptomycin (Sigma–Aldrich, P4333-100ML) in 125 mL shaking flasks (Sigma–Aldrich, CLS431143-50EA) and cultivated in a fed batch cultivation for 13 days. Forty-eight hour after transfection cells were diluted to 400 000 cells mL^{-1} with fresh medium. Successful transfection and protein secretion were determined 6 days after transfection by performing Western Blots. Positive screened samples were initiated to production by the addition of 20% CHO CD EfficientFeed B (Thermo Fisher, A1024001) and a temperature shift to 30 °C. A second feed of 10% was added at day 9 after transfection. The supernatant was clarified by centrifugation and serine-protease inhibitor was then added. For purification, 1 mL of an in house developed anti protein C affinity matrix was used. The harvest sample was filtrated into the matrix and $CaCl_2$ was added to a final concentration of 2 mM. The tube with sample and matrix was then incubated in a cold room overnight. After packing the matrix in a column, it was washed with equilibration buffer (20 mM Tris, 100 mM NaCl, 2 mM CaCl2, pH 7.5) and thereafter a filter was placed on top of the matrix and the column was placed on ASPEC 271 or 274 liquid handlers (Gilson Inc.). After an additional washing step (20 mM Tris, 1 M NaCl, 2 mM $CaCl_2$, pH 7.5) the protein was eluted using a mild elution with EDTA (20 mM Tris, 100 mM NaCl, 2 mM EDTA, pH 7.5) prior a buffer exchange into 1× PBS. After desalting, the protein concentration was determined (Abs). Each purified protein was identified by MS/MS and the purity was analyzed using SDS-PAGE and western blot. Primary Ab for western blotting was a rabbit Anti-C tag polyclonal (GTX18591, Genetex). Glycosylation patterns of the purified proteins were also analyzed using SDS-PAGE.

2.4. Antibodies

Overall, 624 Abs targeting 209 unique secreted human proteins, as well as 11 assay specific controls were included. Majority of the Abs used polyclonal rabbit Abs generated within HPA project (www.proteinatlas.org).[10] The assay specific controls included affinity purified rabbit IgG (P120-301, Bethyl laboratories) in order to control for background binding to rabbit IgG molecules and a blocked bare bead (without coupled Ab) to monitor background binding to the beads. These two will from now on be referred to as "assay controls." In addition, a set of ten monoclonal mouse Abs from BioSystems International,[11] targeting plasma proteins commonly enriched by immuno-capture assays were included.[12] These will be referred to as "internal controls." All Abs used are listed in (Table S1, Supporting Information).

2.5. Coupling of Antibodies to Beads

Bead arrays were created as previously described.[13] Antibodies were diluted to 17.5 µg mL^{-1} in 100 µL 0.1 M 2[N-Morpholino]ethanesulfonic acid (MES)-buffer (M2933, Sigma–Aldrich), pH 4.5, using a pipetting robot (TECAN EVO150), and then coupled to carboxylated color-coded magnetic beads (MagPlex-C, Luminex Corp). In short, 500 000 beads per distinct color-coded region (ID) were located into one wells of 96-well microtiter plates (Greiner BioOne) respectively and washed with 80 µL 0.1 M NaH_2PO_4 (phosphate buffer) pH 6.2 (S3139, Sigma Life Science) with a plate washer/dispenser (EL406, Biotek) on magnet. Subsequently, 50 µL phosphate buffer was added manually. Activation buffer consisting of 10 mg mL^{-1} 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) (C1100, ProteoChem) and 10 mg mL^{-1} sulfo-*N*-hydroxysulfosuccinimide (Sulfo-NHS) (24510, Thermo-Fisher Scientific) in phosphate buffer were subsequently added to the beads, resulting in 0.5 mg EDC and 0.5 mg Sulfo-NHS per well. Activation buffer and beads were incubated for 20 min at 650 rpm at room temperature and washed two times with 100 µL 0.1 M MES. The prediluted Abs were added to the activated beads and incubated for 2 h at 650 rpm at room temperature. After incubation, the Ab-coupled beads were washed three times in 100 µL 1 × PBS (09-9400, Medicago), 0.05% Tween20 (BP337, Fisher Bioreagents) (PBS-T), and re-suspended in 50 µL storage buffer (Blocking Reagent for ELISA, 11 112 589 001, Roche Diagnostics) supplemented with ProClin (4812-U, Sigma–Aldrich). The individual bead IDs were pooled together after overnight blocking at 4 °C, creating six bead stocks containing 65–95 different kinds of Ab-coupled beads, including ten additional control Abs each, coupled to unique bead IDs. The coupling efficiency of the Ab-coupled beads was tested using R-Phycoerythrin-conjugated (RPE) goat anti-rabbit IgG (111-116-144, Jackson ImmunoResearch) and RPE-conjugated goat anti-mouse IgG (115-116-146, Jackson ImmunoResearch). One hundred microliter RPE-conjugated Abs diluted to 0.5 µg mL^{-1} in PBS-T were added to 5 µL Ab-coupled bead stock in different wells, followed by incubation for 20 min at 650 rpm at room temperature. After incubation, wells were washed three times with 100 µL PBS-T before analyzed on a Flexmap 3D instrument (Luminex corp.). Signals for the coupling efficiency were reported in terms of median fluorescence intensities (MFI). Coupled beads were regarded as a failed coupling if the signals obtained were lower than 2 × SD than the mean value for the bead stock. In case of failed coupling, a re-coupling of this specific Ab was performed.

2.6. Biotinylation of Detection Antibodies

Antibodies used as detAbs were biotinylated as described previously.[14] In short, 2 µg of each Ab was diluted in 30 µL PBS-T and then incubated with 5 µL protein A-coated magnetic beads (30 mg mL^{-1}, Dynabeads, 10002D, Invitrogen) for 30 min, room temperature, 650 rpm. After incubation, the Ab-coupled beads were washed three times in 100 µL PBS-T before labeling the Abs with a 150 × molar excess of EZ-Link-NHS-PEG4 – Biotin (21329, Thermo Scientific) dissolved in DMSO (276855, Sigma–Aldrich) for 30 min, room temperature, 650 rpm. The beads were then washed three times in 100 µL PBS-T. The labeled Abs were dissociated from the beads by adding 15 µL 0.2 M acetate (97064-482, VWR), pH 3.2 (elution buffer) to the wells and incubated for 2 min at room temperature, while mixing gently. The supernatants were collected using a magnet and transferred into individual tubes. To buffer the solution, 5 µL of 0.5 M Tris-base (T6066, Sigma–Aldrich), pH 8, was added to each eluate. Subsequently, 5 µL PBS-T were added to each tube and the labeled Abs were stored at 4 °C with an estimated concentration of 0.072 µg μL^{-1}.

The biotinylation efficiency was tested by diluting the labeled Ab to 1 µg mL^{-1} in PBS-T, and adding 25 µL of prediluted labeled Ab to 2 µL of donkey anti-rabbit IgG (711-005-152, Jackson ImmunoResearch) coupled beads followed by incubation for 1 h at 650 rpm at room temperature. After incubation, wells were washed three times with 100 µL PBS-T. Subsequently, 50 µL of a 1:750 dilution of RPE-labeled streptavidin (SA10044, Invitrogen) were added and incubated for 20 min, room temperature, 650 rpm. After incubation, wells were washed three times with 100 µL PBS-T, before analyzed on a Flexmap 3D instrument (Luminex corp.). Abs with signals 50× above background were considered successfully biotinylated.

2.7. Assay Procedure and Read Out

For assay performance, two different batches of commercially available human K2 EDTA mixed gender plasma pool (HMPLEDTA2, Seralab) were serially diluted in plasma dilution buffer to cover a dilution range of 1:4 till 1:3000 in seven steps with equal dilution. The first screening round used a different plasma batch then the rest of the experiments. The plasma dilution buffer consisted of 1 × PBS with 0.5% (w/v) polyvinylalcohol (P8136, Sigma–Aldrich), 0.8% (w/v) polyvinylpyrrolidone (PVP360, Sigma Life Science), 0.1% casein (C5890, Sigma Life Science), and supplemented with 0.5 mg mL^{-1} rabbit IgG. A spike-in serial dilution of standard proteins in plasma dilution buffer was performed, covering a concentration range of 1 µg mL^{-1} to 1 ng mL^{-1}. Blanks of both assay buffer and plasma dilution buffer were added and will be referred to as "blank sample."

The pre-diluted plasma samples and the pre-diluted protein standards (45 µL) were transferred to 5 µL bead stock in an assay plate (Greiner 384-well assay plate) using a liquid handler (SELMA, CyBio) before overnight incubation at 650 rpm at room temperature.

After incubation, the beads were washed three times with 60 µL PBS-T. The biotinylated detAbs were diluted to 1 µg mL^{-1} in PBS-T. Subsequently the beads were incubated for 1.5 h at room temperature at 650 rpm with 25 µL of the respective prediluted detAb. Beads were washed three times with 60 µL PBS-T before incubation with 50 µL of a 1:750 dilution of RPE-labeled streptavidin for 20 min at room temperature at 650 rpm. Finally, beads were washed three times with 60 µL PBS-T, before they were re-suspended in 60 µl PBS-T and analyzed on a Flexmap 3D instrument (Luminex corp.). Binding events were displayed as MFI where at least 50 beads per bead ID were counted.

Each assay plate represents one experimental assay run combined of two 96-well plates containing serial dilutions of human plasma pool as well as two 96-well plates containing serial diluted standard curves of the proteins investigated. Protein and plasma dilution series for the same detAbs were placed on the same 384-plate (Figure S1, Supporting Information). Additionally, interfering plate-effects were avoided by running all measurements for one target protein on the same 384-plate. In total, twenty-six 384-well plates were measured, containing between two and 15 proteins each. A detailed plate layout can be found in Figure S1, Supporting Information.

2.8. Assay Optimization and Validation

The assay design was modified and optimized each time between two phases. After the screening phase, the buffer matrix for the protein standards was changed by adding 1% BSA (A7030, Sigma Life Science) to achieve a higher matrix complexity. Additionally, the length of the dilution curves for the protein standard was extended from a seven-step concentration to a 14-step concentration series in triplicates covering a range of 1 µg mL^{-1} to 1 pg mL^{-1} when evaluating the reproducibility of the assays. The dilution points for plasma were also adapted to the signals achieved during screening, to both cover a broader measuring range, but also to be more suitable for the obtained signals. Thus a seven-step dilution series of human EDTA plasma pool with a consistent step size of three starting between 1:2 and 1:36 in plasma dilution buffer was conducted for each Ab pair. The suspension bead arrays (SBAs) for the selection process were composed with different capAbs for further technical investigations in order to exclude additional off-target interactions. The remaining target proteins were grouped as sets of five into SBAs containing 8–18 Abs, based on their alphabetical order. Each set of SBAs was supplemented with the assay controls, to record possible binding to the beads. For protein quantification and application of the SIA pairs, assays were run on a longitudinal sample set as well as an eight-step protein concentration series, covering the optimized signal range, and measured in triplicates.

2.9. Selection Criteria

After selecting protein targets for the secretome according to availability, and HPA Abs both according to availability and concentration (>0.05 mg mL^{-1} for being chosen either as capAb or detAb), Ab pairs had passed several selection rounds in order to achieve reliably functioning Ab pairs. This process was divided into three phases: an initial screening phase, which was sub-divided into two rounds, a selection phase and an

application phase. For all phases, MFIs were registered for each bead ID and sample.

Annotation after screening phase was performed manually. Hereby, Ab pairs were grouped based on their functionality into four different categories: 1) dilution-dependent curves with protein and plasma, 2) dilution-dependent curves with protein only, 3) dilution-dependent curves with plasma only, and 4) no dilution dependent curves. This was assessed according to the shape and concentration dependency of the curve for the expected pair.

Pairs being processed to be further tested had to be assigned to group 1 as well as reach a maximum signal intensity of at least 150 MFI in order to report only signals above an average background. To limit the number of total assays, one detAb was chosen per protein. For the detAbs, any concentration dependent binding for the other Ab-coupled beads in the SBA, such as the internal controls, were used as exclusion criteria. As additional criteria, a pair was chosen upon showing the widest range of detectable concentrations of proteins in buffer and plasma, an overall lower background level in antigen-free samples, and indications about possible interferences or off-target recognition of other captured proteins. Finally, the available Ab volume was considered for the polyclonal binders.

After an additional testing round of the chosen pairs in triplicates of plasma dilution and protein standard dilution the CV was determined for each dilution point averaging it across all dilution steps to find the best performing Ab pairs (see **Figure 4**A–D). As cut-off criteria a CV of 3.3% in protein assays and 4.3% in plasma assays was defined. In addition, pairs requiring more than 12.5 µL sample (representing a 1:4 sample dilution) or more than 150 ng (=3000 ng mL^{-1} as the highest concentration point) per assay were excluded in respect to sample and protein consumption. For the technical replicates during the selection phase the upper limit of quantification (ULOQ), lower limit of quantification (LLOQ), LOD, and 50% effective dose (ED50) were calculated. One pair per target was prioritized and the calculated ED50 point was chosen as the optimal sample dilution point. In cases of similar performance, Ab pairs generated toward different binding regions were prioritized. Also, Ab pairs with the superior LLOQ were preferred (see Table S2, Supporting Information).

Before processing the remaining Ab pairs into the final application phase and measuring them on a selection of 72 samples from a healthy longitudinal cohort, different Ab pairs were combined into possible duplex combinations with the same determined optimal sample dilution. The final concentration measured for each sample was calculated by transforming the measured MFI signal intensities into a concentration value according to the dilution curve obtained from the 5-parametic fit and multiplying it with the applied dilution factor of the sample (see **Figure 5**A–D). Samples with protein concentration below the calculated LLOQ or above ULOQ were excluded from further analysis.

2.10. Orthogonal Assays

For orthogonal comparison of the targets, the overlap of the chosen 21 targets was correlated with data achieved by two independent experimental setups for the same sample selection: 1) a recently published targeted MS approach[15] and 2) multiplex proximity extension assays (Olink Bioscience, Uppsala Sweden).[16]

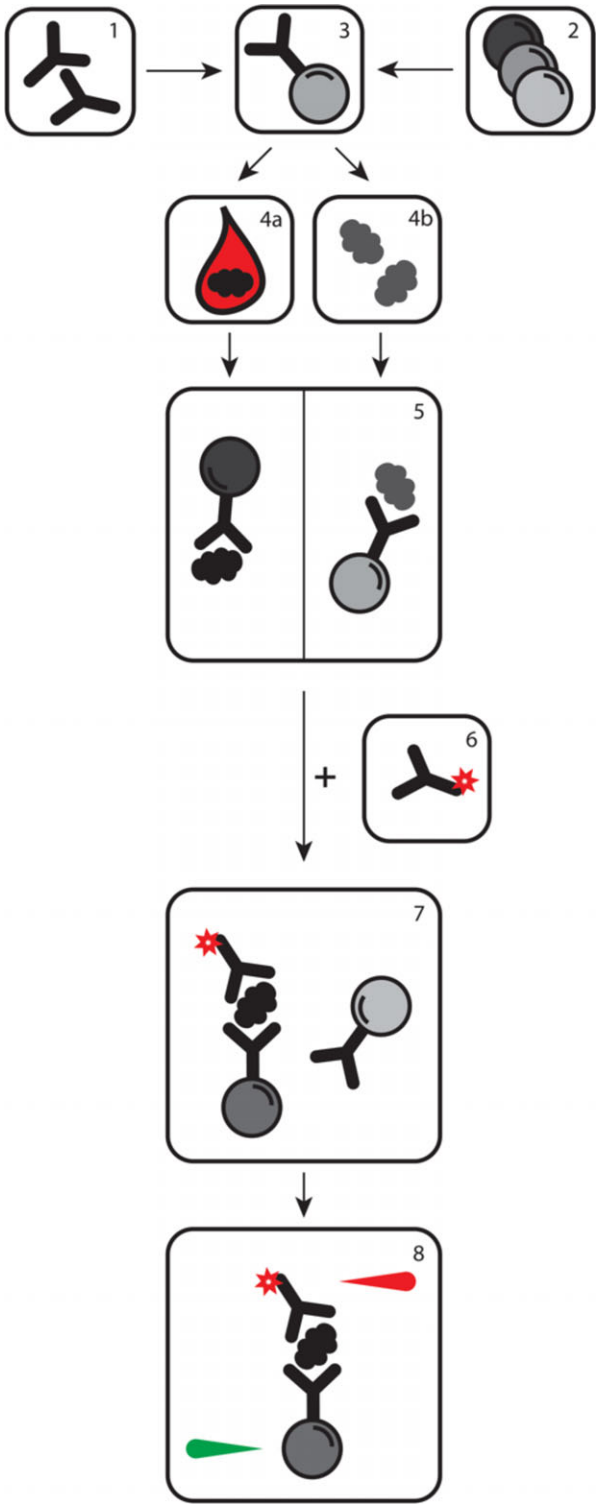
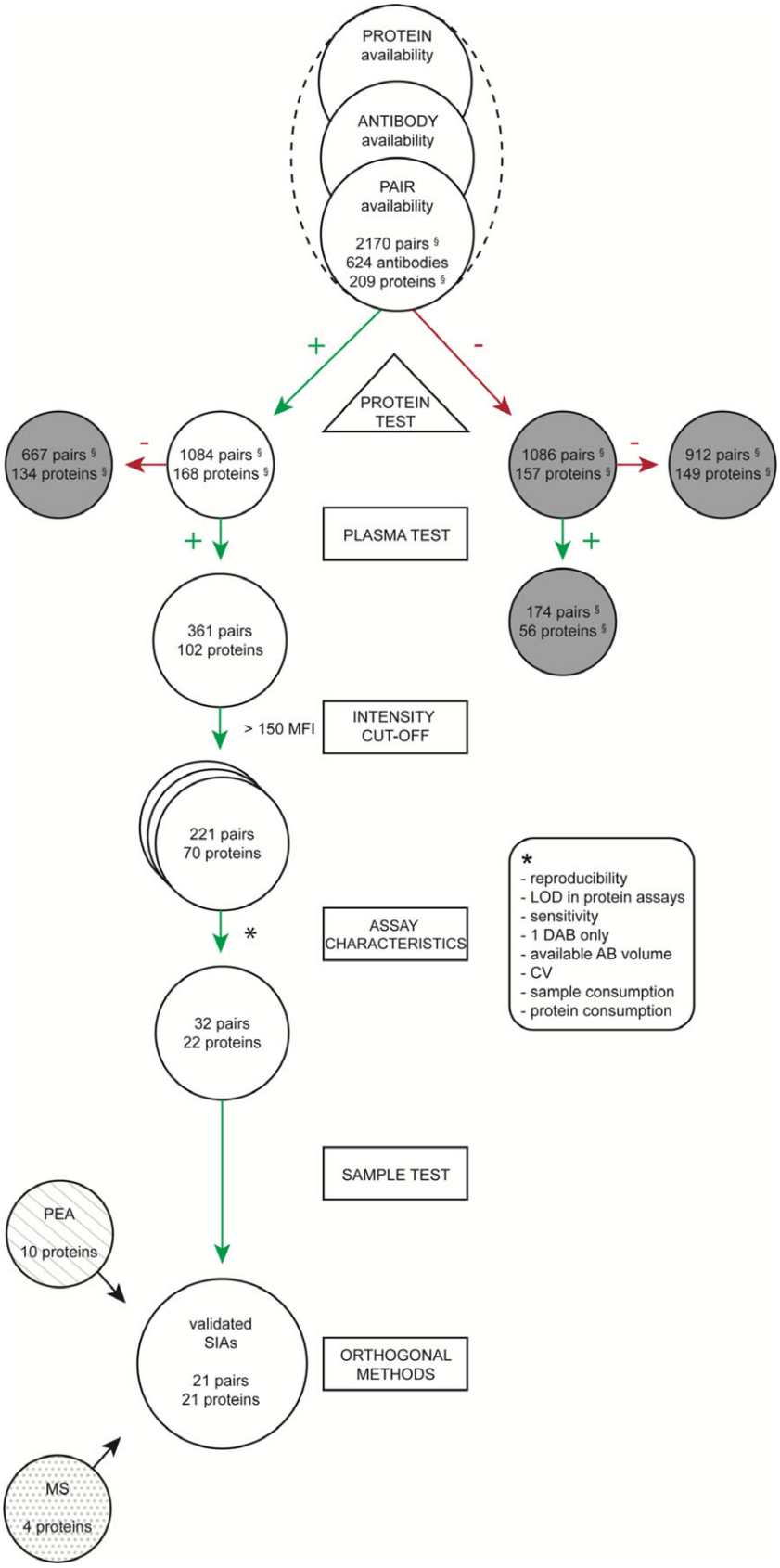


Figure 1. Setup for screening for SIA pairs. For capturing, Abs (1) were immobilized onto magnetic color-coded beads (2) and combined into sets of suspension bead arrays (3). A dilution series of EDTA plasma (4a) and protein standard (4b) was performed. Beads were then combined with either EDTA plasma or protein standard (5). After washing off unbound proteins, the captured proteins were detected via biotinylated Abs (6 and 7). The readout occurred by the addition of a streptavidin-fluorophore and using the Luminex systems (8).



For the targeted MS approach 432 samples were prepared semi-automatically using the Bravo liquid handler and subsequently measured using a combination of Ultimate 3000 binary RS nano LC system (Thermo Scientific) with an EASY-Spray ion source connected to an online Q Exactive HF (Thermo Scientific) MS. All plasma samples were stored lyophilized and resuspended by the autosampler. Sample analysis was performed using a previously developed PRM method. Each full MS scan at 60 000 resolution (AGC target $3e^6$, mass range 350–1600 m/z and injection time 110 ms) was followed by 20 MS/MS scans at 30 000 resolution (AGC target $2e^5$, NCE 27, isolation window 1.5 m/z, and injection time 55 ms), which were defined by a scheduled (2 min windows) PRM isolation list that contained 174 paired light and heavy peptide precursors (n (peptides) $= 87$) from 55 QPrESTs directed toward 52 human proteins. The raw MS-files from all study samples were processed in Skyline (version 3.7) and analyzed in R (version 3.4.1) for protein quantification.

For some of the measured plasma proteins, additional validation was achieved by using multiplex proximity extension assays. Each kit contained a microtiter plate measuring 92 protein biomarkers in up to 90 samples. Each well contained 96 pairs of DNA-labeled Ab probes. Samples were incubated in the presence of proximity Ab pairs tagged with DNA reporter molecules. When the Ab pairs bound to their corresponding antigens, the corresponding DNA tails form an amplicon by proximity extension, which can be quantified by high-throughput real-time PCR.[16,17] To minimize inter- and intra-run variation, the data were normalized using both an internal control (extension control) and an interplate control, and then transformed using a predetermined correction factor. The preprocessed data were provided in the arbitrary unit Normalized Protein eXpression (NPX) on a $\log_2$ scale. A high NPX presented high protein concentration.[17]

2.11. Data Analysis (Data Processing, Classification, and Curve Fitting)

Data analysis and visualizations were performed within R (www.rproject.org, version R 3.5.1).[18] To assess reproducibility for overlapping targets between the two screening rounds, corresponding MFI values were log transformed and correlated using Pearson correlation with R^2 values. To assess the binding region for capAbs and detAbs on the screened proteins, the immunogens aminoacidic sequence for each HPA was mapped to the sequences of the corresponding canonical protein (www.proteinatlas.org). Protein sequences were exported from the Uniprot data base (release 2018_07).[19]

To evaluate the performance of Ab pairs during the selection phase, data were $\log_{10}$ transformed and visualized as dilution curves.

Data were $\log_{10}$ transformed and a five-parametric log-logistic model was applied for the dilution curves in the application phase.[20] LOD levels were calculated as $3 \times$ SD of the blank sample above the average blank signal, LLOQ was defined as $10 \times$ SD of the blank sample above the average blank signal,[21] ULOQ was defined as the averaged signal of the highest protein standard concentration point minus its SD. ED50 was calculated using the drc package.[20] In instances in which the SD was small, leading to negative output of the five-parametric fit for the LOD values, the MFI values for the blank were manually increased by adding 2 AU. No significant effect on the calculated protein concentrations for those targets could be observed.

Assay CVs within the selection phase were calculated between the duplicated dilutions steps of the plasma protein curve, while during the later application phase the assay variance was calculated with the triplicated sample pool. During this phase two additional layers of variance were calculated: variances for each protein between the 18 individuals (per visit), which will be referred to as interindividual CV, as well as the average variance between the 18 subjects over the course of 1 year (four samples per subject), which will be referred to as intra-individual variance. For visualizing different layers of variance (assay variance, interindividual variance, and intraindividual variance), CVs were calculated and ternary plots were generated using the ggtern package.[22]

Euclidian distances for investigating the personal plasma profile differences were calculated using the daisy function in the R package "cluster".[23] Prior to the calculation, the data underwent an outlier removal process, meaning values above ULOQ and below LLOQ as well as NA values were removed from the data set. The data were then scaled before computing pairwise dissimilarities and Euclidean distances. Additionally, Pearson distance for interindividual and intraindividual correlation was calculated, using the R^2 value.

For the correlation plots between different types of data were used: MS data (fmol μL^{-1}) were mean-centered and both MS data as well as SIA data (ng mL^{-1}) were log-transformed with the binary logarithm, while PEA data (NPX values) were used as provided and correlated using Pearson correlation with R^2 values.

3. Results

We aimed to develop a multiplexed workflow (**Figure 1**) to search and select for Ab pairs for the analysis of proteins secreted into

Figure 2. Workflow. A set of 209 protein targets, covered by 624 Abs and resulting in 2170 corresponding Ab pairs were selected and screened on dilution series of both, recombinant proteins as well as an EDTA plasma pool. For some proteins, more than one protein construct was tested, while some Ab pairs were duplicated (§ numbers include those). All pairs were assessed manually for concentration dependent curves in both plasma and protein. 1084 pairs (168 proteins) showed a concentration dependent curve for the protein standard of which 361 unique pairs additionally detected protein in a concentration dependent manner in plasma. For these 361 pairs, corresponding to 102 unique proteins, an additional signal intensity cutoff criteria of $>$ 150 MFI was implemented. Out of the initial 361 pairs, we used 221 for further studies in triplicates, which corresponded to 70 out of the 102 initial proteins. Results of the triplicate measurements were assessed, implementing technical aspects for exclusion, (e.g., assay reproducibility, LOD in protein assays, sensitivity toward the target protein in plasma, CV), but also additional criteria with respect to the available Ab volume, sample, and protein consumption. By this, 32 pairs targeting 22 proteins were left for further analysis. One pair per protein was chosen for validation according to LLOQ, using the ED50 point as the optimal sample dilution point. Finally, the selected 22 pairs were applied as SIAs for the determination of protein levels on a longitudinal plasma sample set ($n = 72$). For 21 pairs, a protein quantification was possible, of which 14 could be compared orthogonally with data from targeted plasma mass spectrometry analysis (MS) or solution-based proximity extension assays (PEA).

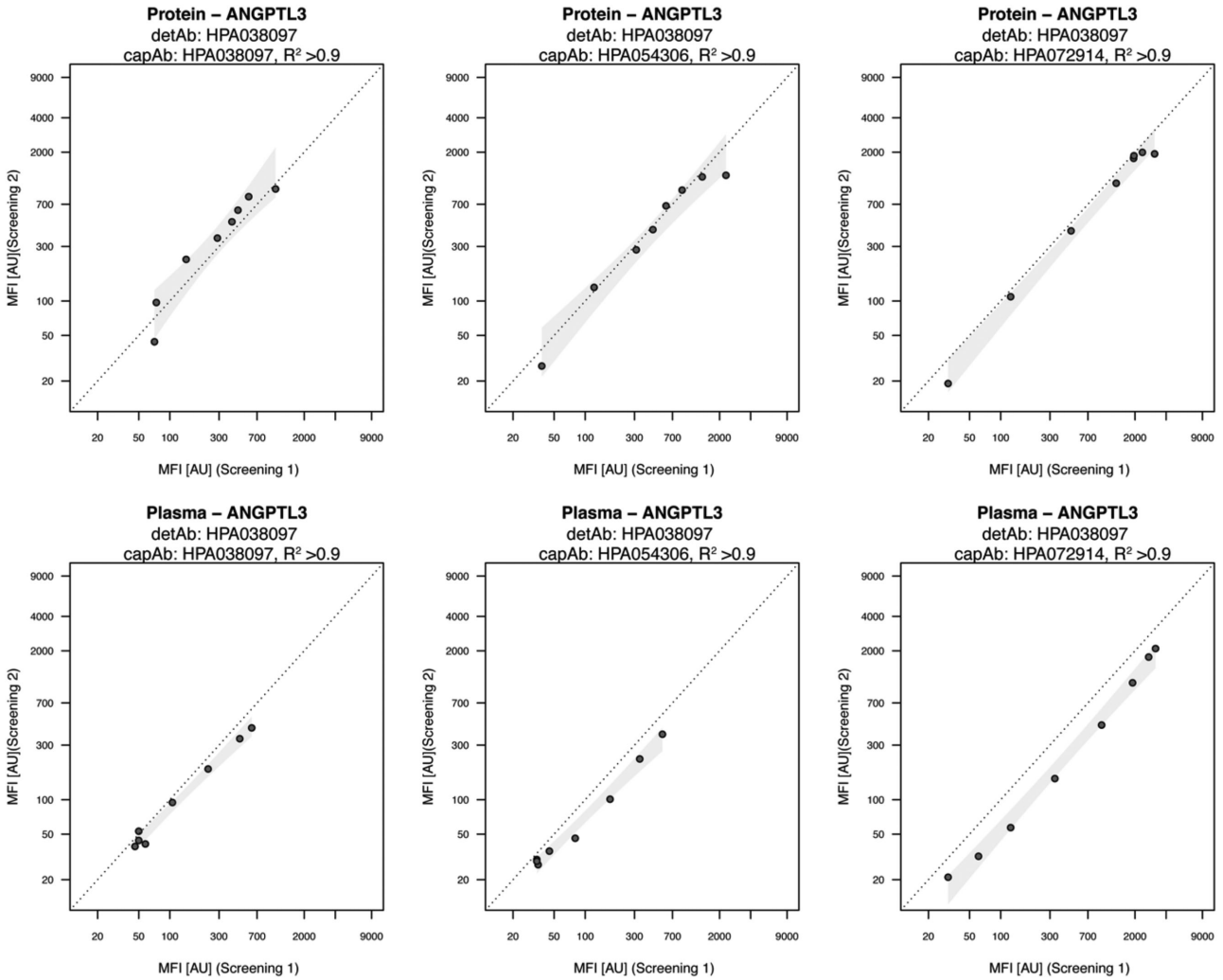


Figure 3. Correlation for ANGPTL3 (detAb: HPA038097) in screening assays. For each detAb, correlation for all capAbs was performed both for protein standard (top) and plasma (bottom) dilution. Correlation between overlapping targets of the two screening rounds was calculated using log10 transformed MFI data and Pearson correlation with R^2 values. Confidence interval for each pair was calculated in R based on a linear model and highlighted in the plots.

Table 1. Annotation of antibody pairs determined during the screening rounds.

	#1	#2	Total
Number of pairs	1118 (100%)	1052 (100%)	1791 (100%)
Passed protein assays only	336 (30%)	331 (31%)	664 (37%)
Passed plasma assays only	77 (7%)	96 (9%)	173 (10%)
Passed both samples types	170 (15%)	247 (23%)	361 (20%)
Failed	535 (48%)	378 (36%)	593 (33%)

human plasma (**Figure 2**). We combined the capabilities of the SBA technology with the resource of HSP's full-length proteins and Abs generated by the HPA project and investigated > 200 proteins as well as ≈1800 possible Ab pairs. The project was designed to be conducted in the following stages: 1) screening for possible Ab pairs in dilution series of protein and plasma samples, 2) preselection of suitable pairs after assessing their apparent functionality, 3) annotation of preselected Ab pairs according to their binding area, 4) selection of Ab pairs for further investigations focusing on technical aspects, 5) preparation of duplex sets for plasma analysis, 6) quantification of plasma protein levels in a longitudinal sample set, and 7) compare these results with data from independent, orthogonal methods.

3.1. Screening for Antibody Pairs

3.1.1. Experimental Study Design

We studied a total of 209 full length proteins and used a pool of EDTA plasma samples to determine and to develop SIAs. The screening was conducted in two rounds of 109 and 124 proteins,

Table 2. Distribution of binding regions.

Capture region	Detection region	Tested Ab pairs	Percentage of tested	Passed Ab pairs	Percentage of passed
N-terminal	N-terminal	176	10%	29	16%
N-terminal	Middle	77	4%	10	13%
N-terminal	C-terminal	104	6%	6	6%
Middle	N-terminal	85	5%	12	14%
Middle	Middle	434	24%	48	11%
Middle	C-terminal	179	10%	21	12%
C-terminal	N-terminal	98	5%	7	7%
C-terminal	Middle	165	9%	29	18%
C-terminal	C-terminal	473	26%	61	13%

where we aimed at replicating the findings from the first round by also including all targets with an apparent functional Ab pair, corresponding to 23 proteins, in screening round 2. For finding Abs from the HPA resource, we chose a concentration cutoff of 0.05 mg mL^{-1} for protein capture, and found 624 Abs for all proteins (see Table S1, Supporting Information). This meant that assays could be developed using an average of three to four Abs per protein, and the coverage ranged from one to eight HPA Abs per protein (see Figure S2, Supporting Information). We chose to combine 49–88 different Abs in one SBA. Among all 624 Abs, we selected those with concentrations $\geq$0.1 mg mL^{-1} and an available volume $\geq$0.5 mL as detection agents. This means that an average of two to three Abs were biotinylated per target protein, and the coverage ranged from one to seven detAbs. Hence, a total of 2170 Ab pairs were investigated, of which 1791 were unique.

The screening rounds were conducted by grouping the proteins into sets of six per assay batch and SBA. Each protein assay contained an SBA of the corresponding Abs as well as those targeting the other five proteins. Each SBA also contained control beads to judge the unspecific binding to the beads, and was distributed into two 384-well plates. The total number of assays per protein was defined by the number of available detAbs, and each detAb was used in eight concentration levels of proteins and eight dilution steps for EDTA plasma. In total, we conducted 9264 assays and generated 553 056 data points for the 209 proteins.

3.1.2. Reproducibility of Screening Results

To assess the reproducibility between the two performed screening rounds, 23 proteins corresponding to 200 unique and target-matched Ab pairs were included in the second screenings. The assay conditions were in terms of dilution factor, number of dilution steps, and the starting concentration for both plasma and protein maintained in the second screening. Out of the 23 targets, assay pairs for 16 targets (70%) revealed reproducible binding curves for both protein and EDTA plasma, as exemplified by ANGPTL3 shown in **Figure 3**. For each target, at least one unique Ab pair had a correlation of $R^2 > 0.92$ for the protein as well as $R^2 > 0.86$ in plasma. In total, 73% of all overlapping pairs had $R^2 > 0.95$ in protein assays, while 59% were > 0.95 in assays with plasma. The pairs toward the additional six targets were regarded as reproduced for the protein assay; however, the detectability in plasma was lower compared to the first screening. A higher degree of reproducibility was observed for targets that provided signal intensity levels further away from the apparent LOD. Additionally, using two different batches of EDTA plasma pool for the two screening rounds could have influenced the reproducibility.

3.2. Preselection of Antibody Pairs

From the generated data, binding curves were manually annotated for plasma and protein in order to classify each Ab pair and select those for further optimization as illustrated in Figure 2. A summary of the outcome of the selection process is shown in **Table 1**. To select pairs based on their apparent functionality, we assessed the shape and concentration dependency of the curve for the expected Ab pair with the used assay conditions. We also considered all other Abs (off-target Ab pairs) included in each SBA as a background measure and noted if unexpected pairs were found in either the protein or plasma samples. All possible pairs were then assigned to one of the following four classifications according to their functionality:

1) Dilution and concentration dependent curves plasma and protein, respectively
2) Concentration dependent curves with protein only
3) Dilution dependent curves with plasma only
4) No dilution or concentration dependent curves

As also summarized in Supplementary Figure 3, from almost 1800 possible Ab pairs there were 170 from screening #1 and 247 from screening #2 that detected their target in plasma and recombinant protein. In total, 361 unique pairs were consequently annotated as "passed" and considered for further assessment and optimization. The remaining pairs did not show a concentration dependent curve for both sample types and may require further time to develop, hence were not considered for further sample analysis.

3.3. Annotation of Preselected Antibody Pairs

As an additional assessment, we investigated the location on the protein to which the pairs of capAb and detAb bound their respective target. We chose to approximate the binding areas

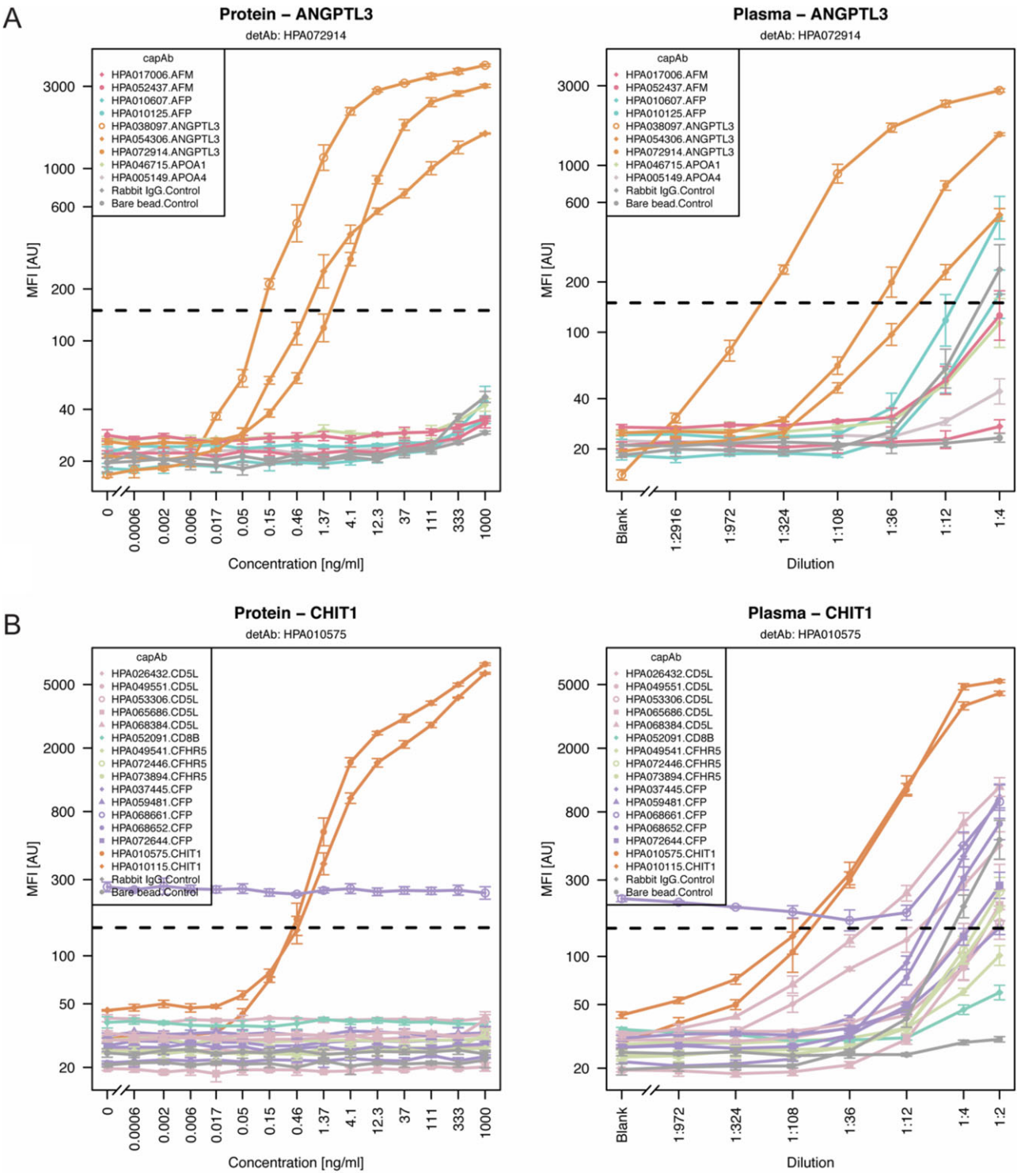


Figure 4. A–D) Dilution curves during selection phase for ANGPTL3, CHIT1, CPA1 and FGF21 in protein standard and plasma. During selection phase, dilution curves for all targets including all capAbs corresponding to one detAb were plotted to evaluate the performance of the different pairs. The signals above background (>150 MFI) are indicated by a horizontal dotted line. Exemplary dilution curves for ANGPTL3 (A), CHIT1 (B), CPA1 (C), and FGF21 (D) in recombinant protein standard (left) and EDTA plasma (right) are shown.

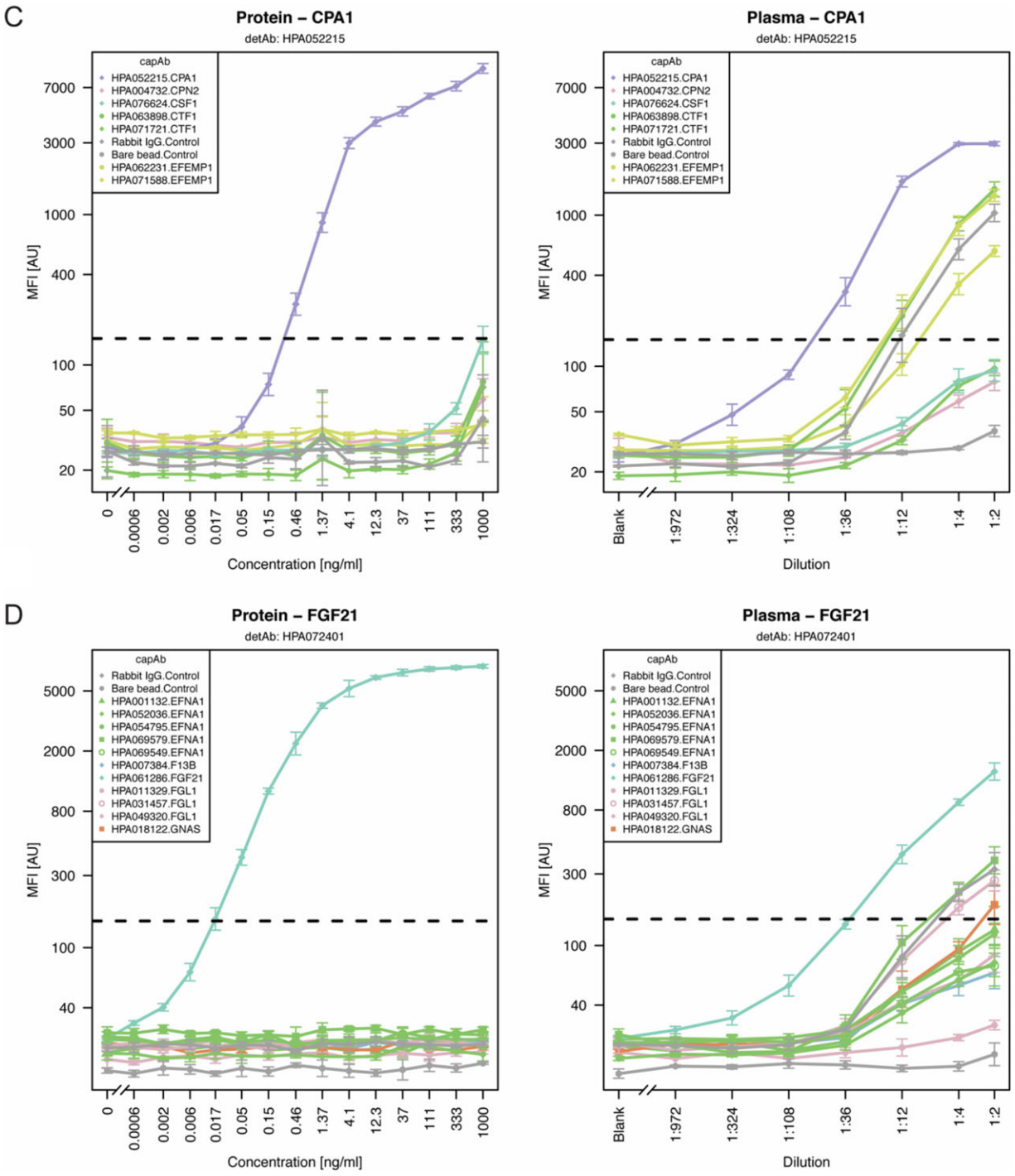


Figure 4. *Continued.*

of the Abs by using their immunogens aminoacidic sequences (22–151 residues in length) and mapping these to the sequences of the canonical protein. Here, we segmented each protein sequence into three equally long parts (N-terminal, middle, and C-terminal). As shown in **Table 2**, there were generally more pairs for constellations that targeted the same region, which was due to using Abs for capture and detection, generated toward an epitope located in the same region. Also, there were more pairs targeting the middle and C-terminal region than in combination with N-terminal binders. Considering the success rate for building SIAs from the screening assessment criteria, we found that on average about 12% of all pre-selected pairs passed these. A slightly higher success rate of 16% was found for purely N-terminal targeted Ab pairs as well as pairs built with a capAb

Table 3. Determination of secreted proteins in plasma ($n = 72$).

Protein	LOD [ng mL^{-1}]	LLOQ [ng mL^{-1}]	ULOQ [ng mL^{-1}]	ED50 [ng mL^{-1}]	Sample dilution	Determined [c] [ng mL^{-1}]	Literature average [c] [ng mL^{-1}]	Commercial SIA kit range [c] [ng mL^{-1}]
AFM	0.2	1.2	1400	89.1	1/400	23 740 ± 7110	75 350[34]	0.078–5
ANGPTL3[b]	0.01	0.05	4	0.5	1/300	149 ± 65	723[28]	0.031–2
APOA4	0.06	1.1	740	173.2	1/3000	32 960 ± 23 170	197 500[34]	0.25–16
CCL16[b]	0.1	0.4	16	2.1	1/30	71 ± 23	11[31]	0.008–0.5
CD5L[a]	0.006	0.19	100	18.6	1/30	441 ± 132	5530[35]	0.156–10
CFHR5[a),b)]	0.1	0.4	73	7.1	1/20	160 ± 37	1660[36]	0.156–10
CFP	1.3	14	1000	234.7	1/10	1520 ± 467	25 000[37]	15.63–1000
CHIT1[a),b)]	0.002	0.01	39	1.2	1/20	65 ± 69	21.4[25]	0.156–10
CPA1[b]	0.00003	0.008	16	0.7	1/30	47 ± 24	908.5[4]	0.65–150
CPN2	0.5	2.7	910	63.1	1/1000	51 840 ± 13 920	55 915.4[4]	1.56–100
EFEMP1	6.6	28	1100	97.2	1/5	381 ± 107	10.5[38]	1.56–100
EFNA1	0.004	0.01	1.2	0.2	1/40	7.8 ± 2.4	7.7[4]	0.156–10
FGF21	0.00002	0.0002	0.4	0.1	1/10	0.3 ± 0.2	0.5[24]	0.031–2
FGL1	0.06	0.2	910	47.3	1/10	613 ± 216	23.3[4]	0.156–10
GNAS	0.003	0.01	3.6	0.7	1/10	2 ± 0.9	5.4[4]	0.312–20
HRG	0.3	1.1	23	6.7	1/1000	15 120 ± 2070	100 000[39]	0.469–30
HSP90B1[b]	0.03	0.1	4.4	1.4	1/300	972 ± 728	464[4]	0.156–10
IGFBP2	0.02	0.2	800	6.3	1/40	379 ± 276	310[31]	0.063–4
INHBC	0.02	0.1	1100	79.0	1/40	386 ± 194	56.2[4]	0.016–1
NPPA	0.003	0.008	24	0.4	1/10	9.1 ± 4.5	0.06[29]	0.016–1
SPON2[a),b)]	0.2	0.6	48	5.5	1/20	68 ± 25	23.6[40]	0.078–5

[a] Measured in serum; [b] Measured in duplex assays.

and detAb targeting the C-terminal and middle, respectively (18%). The lowest success rate of ≈6% was related to preselected binder pairs targeting each one of the termini. Out of the total 1791 pairs, there were 138 pairs passing the selection process of which both Abs targeted the same region.

3.4. Selection of Antibody Pairs

In a third step, we aimed at further shortlisting those preselected Ab pairs. As a primary criterion, the generated level of intensity (reported as MFI) was chosen as an additional cutoff in order to report signals that were five times above an average background level determined by the assay controls (MFI = 30 AU). Those pairs that did not reach a maximal MFI from the protein assay curves of MFI > 150 were therefore excluded. Out of the initial 361 pairs, we used 221 for further studies, which corresponded to 70 of the 102 initial proteins.

The subsequent investigations focused on technical aspects such as assay reproducibility, the apparent LOD using proteins assays, as well as sensitivity of detecting the target protein in plasma samples. All analyses were conducted using triplicates of protein concentration series. To further resemble a sample matrix of higher complexity, the buffer used for technical assessment of protein assays was supplemented by adding 1% BSA. BSA was not added to the buffer for plasma analysis. For each protein, one detAb was prioritized to limit the number of total assays. In cases where several detAbs were available after preselection, additional criteria for prioritization were applied: Ab pairs with the widest range of detectable concentrations of proteins in buffer and plasma, an overall lower background level in antigen-free samples, and no previous indications about possible interferences or offtarget recognition of other captured proteins. The latter was possible to be observed during the screening phase, where five other proteins were also tested in parallel, as each SBA was built with a common set of 49–88 Abs covering six proteins. For the detAbs, any concentration dependent binding for other Ab-coupled beads in the SBA, such as the internal controls, were added as exclusion criteria. Finally, the available Ab volume was considered for the polyclonal binders.

For the selection processes of Ab pairs, target proteins were regrouped into new sets of five protein targets. Concentration of the proteins and the dilution of EDTA plasma were adapted for each individual target according to the data obtained during the screening. Each assay therefore covered a broader range of concentrations in order to determine the optimal dilution point for plasma analysis. Exemplified for SIA pairs targeting ANGPTL3, CHIT1, CPA1, and FGF21, shown in Figure 4A–D, protein detection was specific and accompanied by only very minor increase in signals from other beads. When analyzing plasma, we found that background signals from other Abs arose when using plasma at a lower dilution than 1:50 dilutions. Still, at a plasma dilution of 1:12, the intended signals were fivefold above any other binder pair.

Out of the 221 pairs targeting 70 proteins, we found 43 pairs for 27 proteins suitable for further analysis according to the criteria stated above. To further find the best performing Ab pairs for one protein, we determined the CV by calculating the

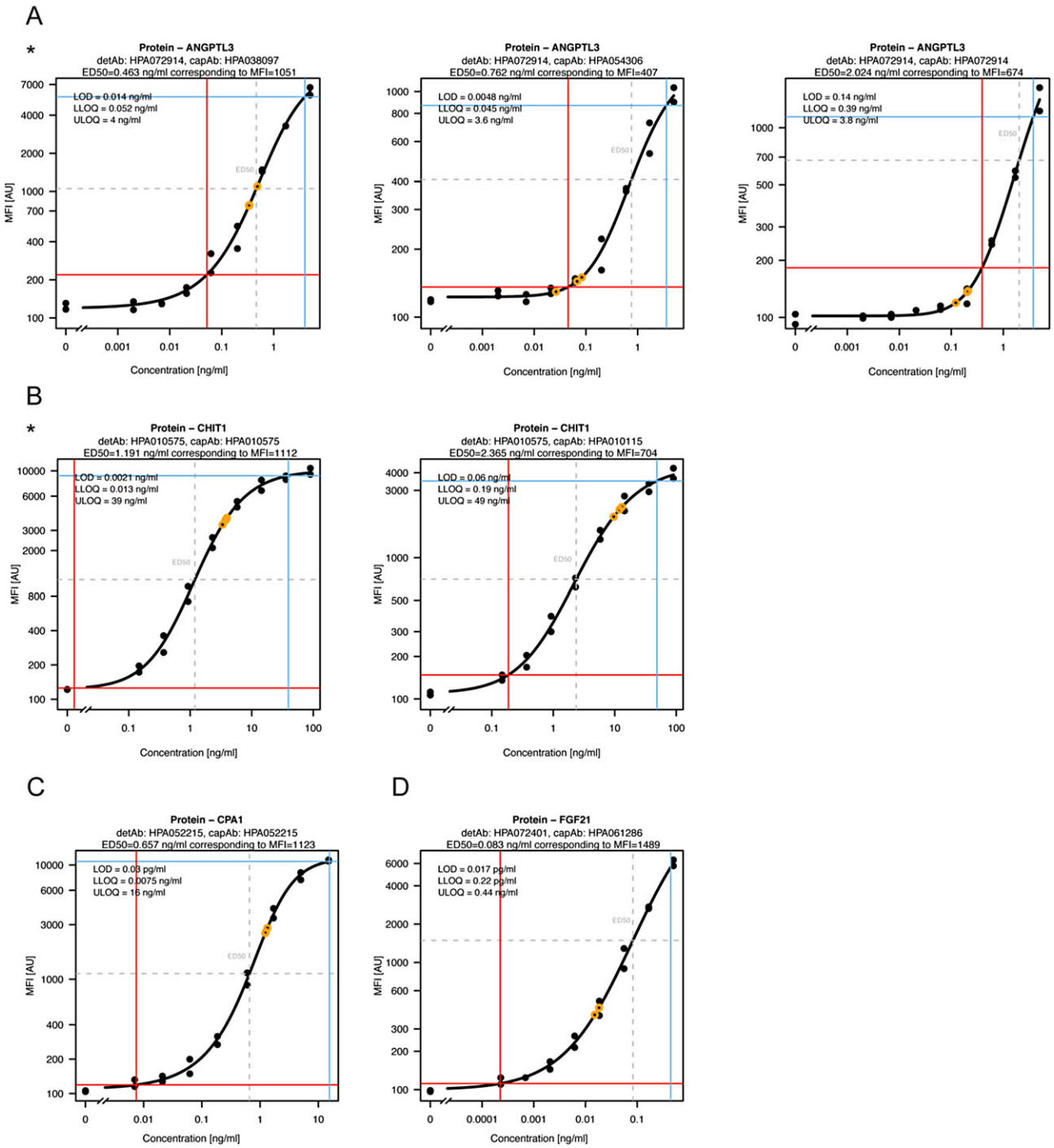


Figure 5. A–D) Protein concentration for ANGPTL3, CHIT1, CPA1, and FGF21 in application phase ($n = 72$). To quantify protein concentration in longitudinal samples during application phase, a 5-parametric log-logistic model was applied for the dilution curves of the protein standard for ANGPTL3 (A), CHIT1 (B), CPA1 (C), and FGF21 (D). Additionally, LOD, LLOQ (red dashed lines), ULOQ (blue dashed lines), and ED50 (grey dashed lines) were calculated. Pooled samples (orange) were plotted onto the curve. If several capAbs were included in the assays, the selected pair is highlighted with "*."

variance for each dilution point using $\log_2$ data, and then using the average across all dilutions within this range for ranking the pairs. As shown in the annotation table (Table S2, Supporting Information), the average CV using log-transformed data was 2.2% for plasma and 2% for protein standard and ranged from 0.4% to 7% in plasma and 0.9% to 4.8% in protein. A set of 10 Ab pairs showing averaged CVs for > 3.3% in protein assays and > 4.3% in plasma assays was excluded from further analysis. Prior to choosing the final set of Ab pairs for plasma profiling, sample and protein consumption was considered. Plasma assays requir-

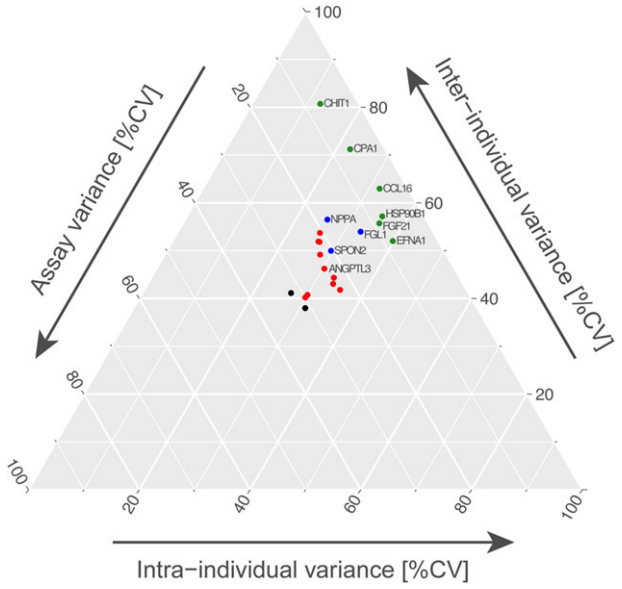


Figure 6. Ternary plot to visualize assay and sample variance. Assay variance was correlated with the inter-individual variance as well as the intra-ndividual variance using a ternary plot. Assay variance was calculated between the triplicated sample pool, inte-individual CV was defined as the variance of each protein between 18 individuals per visit, while the intrain-dividual CV is the average variance between the 18 subjects per target over the course of one year. Proteins showing a low assay variance were highlighted in green (0-10%) and blue (10-20%). Data for ANGPTL3 showed an elevated technical variance, which places this protein among ones highlighted in red (assay variance between 20% and 30%).

ing more samples > 12.5 µL per assay (representing a 1:4 sample dilution) and amounts of proteins exceeding 150 ng (3000 ng mL^{-1} as the highest concentration point) were deprioritized. This led to 32 Ab pairs against 22 proteins for further plasma analysis.

3.5. Preparation of Ab Pairs for Duplexed Plasma Analysis

To achieve a more efficient sample analysis, the data from each protein and plasma dilution curve were compared. The concentration levels for an ED50 were calculated and chosen as the optimal sample dilution point. In order to find the optimal plasma dilution factor per target protein, one Ab pair and ED50 had to be chosen per protein. In cases of similar performance assessment characteristics (see above), Ab pairs generated toward different binding regions were prioritized. Also, Ab pairs with the superior LLOQ were preferred as these generally allow us to cover a broader range of protein levels. To attempt for a higher protein throughput, improve time- and cost-efficiency of the assays, and also reduce sample consumption, we searched for possible combinations of different Ab pairs with the same optimal sample dilution and limited ourselves to assays in duplex. Some combinations were directly excluded due to previously observed incompatibility, so that four duplex assays and 14 single-plex assays remained, as shown in in **Table 3**. We did not find a direct relation between the protein concentration values found in the literature and the degree of sample dilution (Figure S4, Supporting Information).

3.6. Analysis of Protein Levels in a Longitudinal Sample Set

Finally, the selected 22 Ab pairs were used in SIAs for the determination of protein levels in a collection of longitudinal plasma samples. The study set was built of 18 individuals that donated plasma every third month over 1 year. Using the 72 samples collected from four visits, each subject allowed us on the one hand to determine the technical suitability of the selected Ab pairs for analysis of proteins and on the other hand, we could illustrate how protein levels of individuals vary longitudinally and between sample collections.

We quantified 21 of the tested 22 proteins and listed the performance of the assays in Table 3, where the stated protein concentration for each target was calculated from the average concentration over all samples per donor. The protein levels determined here generally agreed well with those found in the literature (see Figure S5, Supporting Information). The standard curves from the new assays are shown in Figure 5A–D and relate to those introduced in Figure 4A–D.

In **Figure 6**, we further compared three layers of variance: technical precision, interindividual differences, as well as longitudinal changes. The ternary plot showed that several proteins, such as CHIT1, CPA1, or FGF21, were stable over time and could be accurately measured, while differing in levels between the donors. The data for ANGPTL3, however, were less conclusive due to an elevated technical variance (CV > 21%, using raw data).

In addition to variance analysis, we used distances from clustering analysis to compare the interindividual differences and the intra-individual differences. Our analysis reveals an average intraindividual Euclidean distance of 3.4 compared to 5.4 between the individuals. This is in line with other observations that protein levels in plasma remain constant over the course of one year and that each person has a unique profile.

3.7. Comparison to Orthogonal Plasma Assays

Finally, we aimed to confirm the data obtained by the selected Ab pairs though using additional analyses. This assessment was based on comparing our data with results from targeted plasma MS analysis[15] and solution-based proximity extension assays (PEA).[16] For above methods, we obtained data sets generated in previous studies of the longitudinal sample analyzed in the application phase (Fagerberg et al., unpublished). Using direct correlation analysis as a proxy to determine the similarity between the generated data sets, protein levels from 14 targets were studied. Of the alternative methods, data for ten proteins only were available for PEA and for four proteins from MS only. As shown in **Table 4** and **Figure 7**A–D, correlations between our protein levels and another affinity-based method, PEA, reached $R^2 = 0.6 \pm 0.2$ while correlations with peptide abundance from MS were $R^2 = 0.3 \pm 0.2$. This illustrates that it was possible to obtain supportive evidence for some of the target proteins, but differences between the assay types in terms of sensitivity and assay interference between the technologies may have contributed to a reduction in concordance. It is worth noting that the assays were performed in different labs and at different timepoints, too. When choosing other capAbs of the SBAs than those shortlisted for the preferred pairs, an additional set of ten capAbs were available to compare

Table 4. Validation of antibody pair with orthogonal methods (PEA, MS).

Protein	Assay 1 (SIA pair of interest) (capAb-detAb)	Assay 2	Pearson's R^2
AFM	HPA052437-HPA017006	HPA017006-HPA017006	0.93
	HPA052437-HPA017006	MS	0.52
ANGPTL3	HPA038097-HPA072914	HPA054306-HPA072914	0.50
	HPA038097-HPA072914	PEA	0.46
APOA4	HPA005149-HPA005149	MS	0.52
CCL16	HPA068091-HPA042909	HPA042909-HPA042909	0.81
	HPA068091-HPA042909	HPA051577-HPA042909	0.94
	HPA068091-HPA042909	PEA	0.63
CD5L	HPA026432-HPA026432	MS	0.16
CFHR5	HPA072446-HPA072446	PEA	0.79
CHIT1	HPA010575-HPA010575	HPA010115-HPA010575	0.97
	HPA010575-HPA010575	PEA	0.85
CPA1	HPA052215-HPA052215	PEA	0.85
CPN2	HPA004732-HPA004732	MS	0.12
EFEMP1	HPA071588-HPA070841	HPA062231-HPA070841	0.77
	HPA071588-HPA070841	PEA	0.13
FGF21	HPA061286-HPA072401	PEA	0.73
IGFBP2	HPA077723-HPA004754	HPA004754-HPA004754	0.89
	HPA077723-HPA004754	PEA	0.60
INHBC	HPA071895-HPA050755	HPA057468-HPA050755	0.93
	HPA071895-HPA050755	HPA050755-HPA050755	0.96
	HPA071895-HPA050755	HPA020729-HPA050755	0.88
	HPA071895-HPA050755	PEA	0.35
SPON2	HPA040170-HPA066095	PEA	0.36
	Available orthogonal data	*n*	Average $R^2 \pm$ SD
	SIA (intra)	10	0.9 ± 0.1
	PEA	10	0.6 ± 0.2
	MS	4	0.3 ± 0.2

the data from the primary Ab pairs with. Since the data from these assays were obtained from the same sample incubation and used the same detAb, it was less surprising but reassuring to find a high correlation between the primary and additional Ab pairs of $R^2 = 0.9 \pm 0.1$ (Table 4).

4. Discussion

This study describes a workflow based on a multiplexed bead-based platform and high-quality reagent resources to systematically screen, select, and apply pairs of Abs for the quantification of secreted proteins in human plasma. Starting from 1791 unique Ab pairs built on 624 unique Abs, two rounds of screening were conducted on parallel dilution series of EDTA plasma and the recombinant full-length target protein. We found that 20% or 361 of all possible pairs detected the recombinant as well as the plasma protein in a concentration dependent manner. We applied a selected set of 32 SIA pairs to study the protein levels in plasma collected from 18 subjects every third month over 1 year and finally confirmed these findings by using orthogonal assays for 14 targets. For six of these, the Pearson correlation between the orthogonal assay and the validated SIA pair was $R^2 > 0.6$.

Our study was based on the use of available HSP proteins and HPA Abs. The polyclonal binders undergo a stringent quality assessment for the use of immunoassays such as Western blot, immunohistochemistry, and confocal microscopy. However, the functionality in other types of assays and samples require separate efforts and we did not find a direct link between building functional SIA pairs with pre-assessing these binders in Western blots on plasma (data not shown). Further to this, we did not include affinity reagents generated by other providers, which may have limited us in providing larger number of assays in the end. In addition, we estimate that many of the capAbs that are currently part of "nonfunctional" pairs could indeed enrich the target protein of interest, but the tested detAb was not suitable in combination with these. We also acknowledge the fact that polyclonal Abs need to be regenerated for extended use and are therefore less suitable for clinical utility. However, applying stringent validation criteria and generation of larger batches may still open up these binders for large-scale studies and exploratory research. In addition, identifying suitable antigens from studies based on polyclonal Abs may streamline the development of monoclonal and recombinant binder libraries.

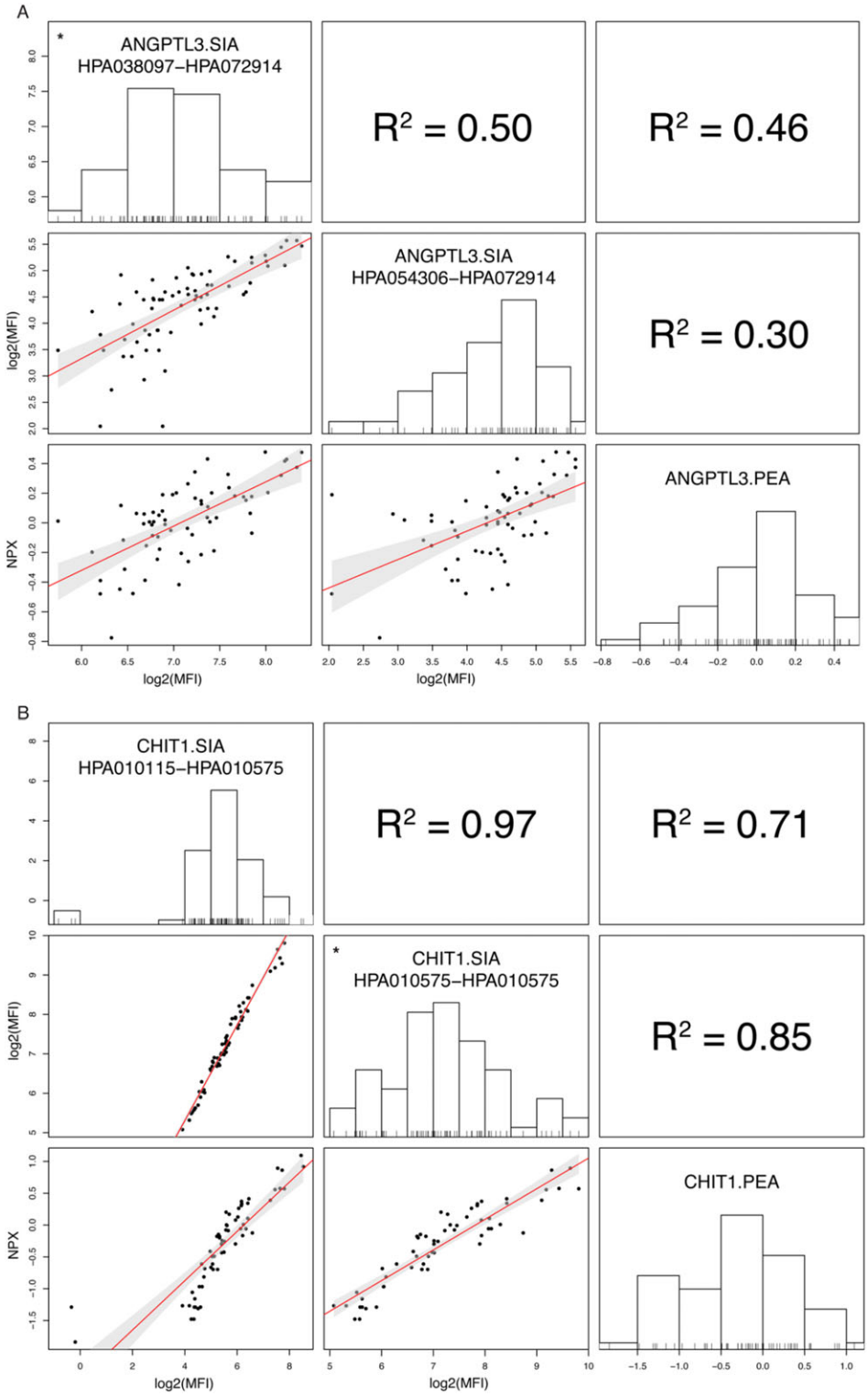


Figure 7. A–D) Comparison of assays for detecting in human plasma. Protein concentration achieved from the SIA pairs was compared with solution-based proximity extension assays. The paired plots for ANGPTL3, CHIT1, CPA1, and FGF21 show the distribution of the data as histograms for each assay, the R^2 from the Pearson correlation between different SIA pairs and PEA assays. For visualization and correlation, the SIA data was log2-transformed and NPX values from PEA were used. The Abs used for the SIA pairs are described as capture-detection, and the numbers stated in the upper right-hand corner refer to the R^2 correlation values: For ANGPTL3 (A) and CHIT1 (B) data from two SIA pairs were compared to each other ($R^2 = 0.50$ or 0.97) and to data from PEA ($R^2 = 0.46$ and 0.30 for ANGPTL3; $R^2 = 0.71$ and 0.85 for CHIT1). For CPA1 (C) and FGF21 (D), data from one SIA pair were compared to data from PEA ($R^2 = 0.85$ or 0.73). The validated SIA pairs that were included to profile the longitudinal samples are highlighted with "*."

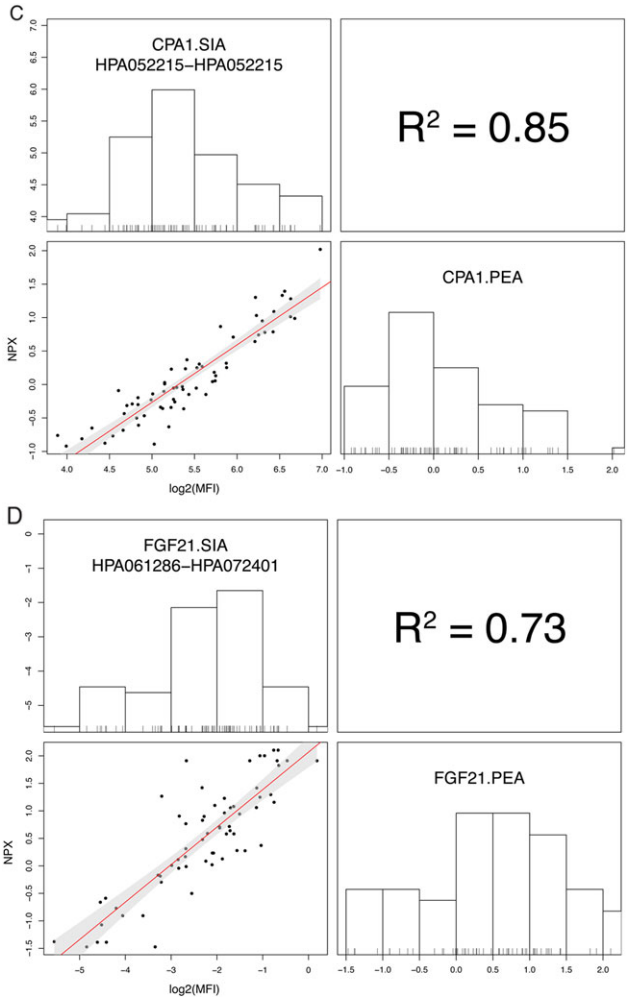


Figure 7. *Continued.*

The proteins that we quantified here are important indicators of health status relevant for different diseases: FGF21 for example is a known key regulator in lipid and glucose metabolism, which is increased in conditions such as type 2 diabetes, obesity, and nonalcoholic fatty liver disease.[24] CHIT1 serves as a neuroinflammatory marker and shows increased concentrations in sALS (sporadic amyotrophic lateral sclerosis).[25] ER stress is hypothesized to lead to hereditary pancreatitis may promote the development of pancreatic ductal adenocarcinoma. CPA1 is among the highest expressed genes in acinar cells, thus CPA1 proteins are expected to cause more ER stress than lower expressed pancreatic enzymes and are indicated to be associated not only with pancreatic cancer development, but also its susceptibility.[26] ANGPTL3 is a novel factor modulating the plasma lipoprotein metabolism[27] and may additionally contribute to uremic dyslipidemia.[28] Plasma levels of NPPA have been described as prognostic predictors in patients with chronic heart failure, but are also known to reflect the severity of left ventricular hemodynamic dysfunction.[29] It has been suggested that FGL1 plays a role in liver protection and liver regeneration, but it also has the potential to serve as a target for the treatment of gastric cancer and to predict gastric cancer prognosis.[30] In addition, CCL16, a human CC chemokine, has been shown to be differentially expressed in ovarian cancer.[31,32]

The study presented here uses pooled plasma collected from nondiseased subjects. This sample source may have limited the possibility to detect those proteins that increase with inflammation, infection, or other diseases. In addition, more assay optimization may have been necessary to rescue some of those pairs that detected their target protein but did not reveal signals above background in plasma. It may further be necessary to choose other, more sensitive detection systems and thereby sacrificing some of the SBA's capabilities in terms of target and sample throughput. Some Abs showed binding to other than their intended targets once higher amounts of plasma (in particular 1:2). These points at further optimization of the assays are needed in terms of blocking agents and that not all Ab pairs are compatible with another. Additionally, the apparent LOD and LLOQ levels may increase once more complex buffer solutions are used.

To our current knowledge, this is one of the largest and first systematic study to screen for SIAs. We focused on the plasma secretome, as plasma is an important sample for clinical routine

analysis and a highly interesting source for searching for disease related proteins. We assumed that the proteins secreted into the blood stream remain detectable in solution when plasma is being prepared. In comparison to studying proteins in cell lysates, proteins found in plasma do not require to be extracted, hence the need to apply strong detergents can be omitted. However, we acknowledge that proteins may precipitate or denature during sample processing. Some proteins are known to be unstable and degrade over time, hence the possibility to detect these decreases with the age of the sample.

In addition to expand our possibilities to measure proteins actively secreted into blood, highly multiplexed immunoassays as well as MS-based approaches require more targeted assays to validate and quantify potential findings in larger number of samples as well as using orthogonal methods. One advantage of our SIA screening approach is that it provides a foundation for the development of more sensitive assays in human plasma, and it enables to build quantitative assays clinical routines, and allows the transfer of these validated pairs of Abs onto other immunoassay platforms or emerging technologies. It holds a great value to expand this list of assays for the plasma secretome to measure and quantify the plasma components of interest on a protein level. Considering the targets included in our study, this suggests that the detectability of proteins in plasma is predominantly depended on the interplay between the available reagents, the technology, and protein stability. We are, however, well aware that many other, such as the low abundant cytokines,[33] require minimal sample dilutions to detect proteins of pg mL^{-1} concentrations, and we suggest to include necessary controls and considerate elevated and unspecific background binding.

In summary, multiplexed bead arrays were used to screen for functional Ab pairs in proteins and plasma samples. With a success rate of 20%, we found that investing at least three different Abs per target protein and assessing different capture–detection combinations was necessary to obtain Ab pairs for protein quantification. While further assay optimization, additional Abs, and target-centric studies will be needed to assess the utility of these Ab pairs, we could not observe a trend toward any binding-site preference for building functional assays for secreted proteins. Considering the need to generate renewable reagents for extended use, our study provides valuable leads on selecting and building Ab pairs even with other reagents than those used here.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

R.S.H., A.B., M.J.I., L.S., M.D., E.B., S.B., and C.F. planned and performed immunoassays. R.S.H., A.B., U.Q., and J.M.S. analyzed data. F.E. performed MS assays and analysis. T.D.C. analyzed data. L.F. provided Olink data. J.R. and H.T. supervised the protein production and H.T. provided the project with full-length proteins. M.U. and J.M.S. conceived the study. U.Q. and J.M.S. supervised the study. R.S.H., A.B., and J.M.S. wrote the manuscript with scientific input from all coauthors. The authors thank everyone at the Affinity Proteomics and Clinically Applied Proteomics groups at SciLifeLab in Stockholm for their continuous fruitful discussion, access to instrumentation, and input to the presented work. The authors also thank everyone at the Human Protein Atlas (HPA) and at the Human Secretome Project (HSP) for their support. The KTH Center for Applied Precision Medicine (KCAP) funded by the Erling-Persson Family Foundation and is acknowledged for financial support. This work was supported by grants from Science for Life Laboratory and the Knut and Alice Wallenberg Foundation is acknowledged for funding the HPA project and the HSP. HSP was also funded by AstraZeneca and Novo Nordisk Foundation. The work leading to this publication has received support from the Innovative Medicines Initiative Joint under grant agreement no. 115317 (DIRECT), resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies' in-kind contribution.

Conflict of Interest

U.Q. is employee of Atlas Antibodies AB. M.U. is co-founder of Atlas Antibodies AB. F.E., L.F., J.R., H.T., and J.M.S. acknowledge formal links to Atlas Antibodies AB.

Keywords

antibodies, plasma, sandwich assays, screening, secreted proteins

Received: January 8, 2019
Revised: June 17, 2019
Published online: July 22, 2019

[1] B. Ayoglu, A. Haggmark, M. Neiman, U. Igel, M. Uhlen, J. M. Schwenk, P. Nilsson, *Expert Rev. Mol. Diagn.* **2011**, *11*, 219.
[2] M. Uhlen, L. Fagerberg, B. M. Hallstrom, C. Lindskog, P. Oksvold, A. Mardinoglu, A. Sivertsson, C. Kampf, E. Sjostedt, A. Asplund, I. Olsson, K. Edlund, E. Lundberg, S. Navani, C. A. Szigyarto, J. Odeberg, D. Djureinovic, J. O. Takanen, S. Hober, T. Alm, P. H. Edqvist, H. Berling, H. Tegel, J. Mulder, J. Rockberg, P. Nilsson, J. M. Schwenk, M. Hamsten, K. von Feilitzen, M. Forsberg, L. Persson, F. Johansson, M. Zwahlen, G. von Heijne, J. Nielsen, F. Ponten, *Science* **2015**, *347*, 1260419.
[3] M. Uhlen, H. Tegel, Å. Sivertsson, C.-C. Kuo, J. M. Gutierrez, N. E. Lewis, B. Forsström, M. Dannemeyer, L. Fagerberg, M. Malm, H. Vunk, F. Edfors, A. Hober, E. Sjöstedt, D. Kotol, J. Mulder, A. Mardinoglu, J. M. Schwenk, P. Nilsson, M. Zwahlen, J. Ottosson Takanen, K. von Feilitzen, C. Stadler, C. Lindskog, F. Pontén, J. Nielsen, B. O. Palsson, A.-L. Volk, M. Lundqvist, A. Berling, A.-S. Svensson, S. Kanje, H. Enstedt, D. Afshari, S. Ekblad, J. Scheffel, B. Katona, J. Vuu, E. Lindström, L. Xu, R. Mihai, L. Bremer, M. Westin, M. Muse, L. M. Mayr, S. Knight, S. Göpel, R. Davies, P. Varley, D. Hatton, R. Fields, B. G. Voldborg, J. Rockberg, L. Holmberg Schiavone, S. Hober, *bioRxiv* **2018**.
[4] J. M. Schwenk, G. S. Omenn, Z. Sun, D. S. Campbell, M. S. Baker, C. M. Overall, R. Aebersold, R. L. Moritz, E. W. Deutsch, *J. Proteome Res.* **2017**, *16*, 4299.
[5] U. Landegren, J. Vanelid, M. Hammond, R. Y. Nong, D. Wu, E. Ulleras, M. Kamali-Moghaddam, *Anal. Chem.* **2012**, *84*, 1824.
[6] P. E. Geyer, N. J. W. Albrechtsen, S. Tyanova, N. Grassl, E. W. Iepsen, J. Lundgren, S. Madsbad, J. J. Holst, S. S. Torekov, M. Mann, *Mol. Syst. Biol.* 901, **2016**, *12*.
[7] V. Emilsson, M. Ilkov, J. R. Lamb, N. Finkel, E. F. Gudmundsson, R. Pitts, H. Hoover, V. Gudmundsdottir, S. R. Horman, T. Aspelund, L. Shu, V. Trifonov, S. Sigurdsson, A. Manolescu, J. Zhu, O. Olafsson, J. Jakobsdottir, S. A. Lesley, J. To, J. Zhang, T. B. Harris, L. J. Launer, B.

Zhang, G. Eiriksdottir, X. Yang, A. P. Orth, L. L. Jennings, V. Gudnason, *Science* **2018**, *361*, 769.
[8] M. Del Campo, W. Jongbloed, H. A. Twaalfhoven, R. Veerhuis, M. A. Blankenstein, C. E. Teunissen, *Front Neurol.* **2015**, *6*, 202.
[9] G. Bergström, G. Berglund, A. Blomberg, J. Brandberg, G. Engström, J. Engvall, M. Eriksson, U. de Faire, A. Flinck, M. G. Hansson, B. Hedblad, O. Hjelmgren, C. Janson, T. Jernberg, Å. Johnsson, L. Johansson, L. Lind, C. G. Löfdahl, O. Melander, C. J. Östgren, A. Persson, M. Persson, A. Sandström, C. Schmidt, S. Söderberg, J. Sundström, K. Toren, A. Waldenström, H. Wedel, J. Vikgren, B. Fagerberg, A. Rosengren, *J. Intern. Med.* **2015**, *278*, 645.
[10] P. Nilsson, L. Paavilainen, K. Larsson, J. Odling, M. Sundberg, A. C. Andersson, C. Kampf, A. Persson, C. Al-Khalili Szigyarto, J. Ottosson, E. Björling, S. Hober, H. Wernérus, K. Wester, F. Pontén, M. Uhlen, *Proteomics* **2005**, *5*, 4327.
[11] M. Guergova-Kuras, I. Kurucz, W. Hempel, N. Tardieu, J. Kadas, C. Malderez-Bloes, A. Jullien, Y. Kieffer, M. Hincapie, A. Guttman, E. Csanky, B. Dezso, B. L. Karger, L. Takacs, *Mol. Cell. Proteomics* **2011**, *10*. M111.010298.
[12] C. Fredolini, S. Bystrom, L. Sanchez-Rivera, M. Ioannou, D. Tamburro, F. Ponten, R. M. Branca, P. Nilsson, J. Lehtio, J. M. Schwenk, *Sci. Rep.* **2019**, *9*, 8324.
[13] K. Drobin, P. Nilsson, J. M. Schwenk, *Methods Mol. Bio.l* **2013**, *1023*, 137.
[14] M. Dezfouli, S. Vickovic, M. J. Iglesias, P. Nilsson, J. M. Schwenk, A. Ahmadian, *Proteomics* **2014**, *14*, 14.
[15] F. Edfors, B. Forsstrom, H. Vunk, D. Kotol, C. Fredolini, G. Maddalo, A. S. Svensson, T. Bostrom, H. Tegel, P. Nilsson, J. M. Schwenk, M. Uhlen, *J Proteome Res* **2019**, *18*, 2706.
[16] M. Lundberg, A. Eriksson, B. Tran, E. Assarsson, S. Fredriksson, *Nucleic. Acids. Res.* **2011**, *39*, e102.
[17] E. Assarsson, M. Lundberg, G. Holmquist, J. Bjorkesten, S. B. Thorsen, D. Ekman, A. Eriksson, E. Rennel Dickens, S. Ohlsson, G. Edfeldt, A. C. Andersson, P. Lindstedt, J. Stenvang, M. Gullberg, S. Fredriksson, *PLoS One* **2014**, *9*, e95192.
[18] R Core Team, *R A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria **2018**.
[19] T. UniProt Consortium, *Nucleic. Acids. Res.* **2018**, *46*, 2699.
[20] C. Ritz, F. Baty, J. C. Streibig, D. Gerhard, *PLoS One* **2015**, *10*, e0146021.
[21] D. A. Armbruster, T. Pry, *Clin Biochem Rev* **2008**, *29*, S49.
[22] N. Hamilton, M. Ferry, *J. Stat. Soft.* **2018**, *87*, 1.
[23] a) L. Kaufman, P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*, Wiley, Hoboken, NJ, **2009**; b) M. Maechler, P. Rousseeuw, A. Struyf, M. Hubert, K. Hornik, **2018**, unpublished.
[24] K. Gariani, G. Drifte, I. Dunn-Siegrist, J. Pugin, F. R. Jornayvaz, *Endocr. Connect.* **2013**, *2*, 146.
[25] P. Oeckl, P. Weydt, P. Steinacker, S. Anderl-Straub, F. Nordin, A. E. Volk, J. Diehl-Schmid, P. M. Andersen, J. Kornhuber, A. Danek, K. Fassbender, K. Fliessbach, D. German Consortium for Frontotemporal Lobar, H. Jahn, M. Lauer, K. Muller, A. Knehr, J. Prudlo, A. Schneider, D. R. Thal, D. Yilmazer-Hanke, J. H. Weishaupt, A. C. Ludolph, M. Otto, *J Neurol Neurosurg. Psychiatry* **2019**, *90*, 4.
[26] K. Tamura, J. Yu, T. Hata, M. Suenaga, K. Shindo, T. Abe, A. MacGregor-Das, M. Borges, C. L. Wolfgang, M. J. Weiss, J. He, M. I. Canto, G. M. Petersen, S. Gallinger, S. Syngel, R. E. Brand, A. Rustgi, S. H. Olson, E. Stoffel, M. L. Cote, G. Zogopoulos, J. B. Potash, F. S. Goes, R. W. McCombie, P. P. Zandi, M. Pirooznia, M. Kramer, J. Parla, J. R. Eshleman, N. J. Roberts, R. H. Hruban, A. P. Klein, M. Goggins, *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 4767.
[27] R. Koishi, Y. Ando, M. Ono, M. Shimamura, H. Yasumo, T. Fujiwara, H. Horikoshi, H. Furukawa, *Nat. Genet.* **2002**, *30*, 151.
[28] T. Shoji, S. Hatsuda, S. Tsuchikura, E. Kimoto, R. Kakiya, H. Tahara, H. Koyama, M. Emoto, T. Tabata, Y. Nishizawa, *Atherosclerosis* **2009**, *207*, 579.
[29] T. Goto, H. Takase, T. Toriyama, T. Sugiura, Y. Kurita, N. Tsuru, H. Masuda, K. Hayashi, R. Ueda, Y. Dohi, *Nephron* **2002**, *92*, 610.
[30] Y. Zhang, H. X. Qiao, Y. T. Zhou, L. Hong, J. H. Chen, *Mol Med Rep* **2018**, *18*, 1465.
[31] M. Polanski, N. L. Anderson, *Biomark Insights* **2007**, *1*, 1.
[32] G. Mor, I. Visintin, Y. Lai, H. Zhao, P. Schwartz, T. Rutherford, L. Yue, P. Bray-Ward, D. C. Ward, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 7677.
[33] S.-B. Ahn, A. Khan, *EuPA Open Proteomics* **2014**, *3*, 78.
[34] H. Dieplinger, D. P. Ankerst, A. Burges, M. Lenhard, A. Lingenhel, L. Fineder, H. Buchner, P. Stieber, *Cancer Epidemiol. Biomarkers Prev.* **2009**, *18*, 1127.
[35] T. Yamazaki, M. Mori, S. Arai, R. Tateishi, M. Abe, M. Ban, A. Nishijima, M. Maeda, T. Asano, T. Kai, K. Izumino, J. Takahashi, K. Aoyama, S. Harada, T. Takebayashi, T. Gunji, S. Ohnishi, S. Seto, Y. Yoshida, Y. Hiasa, K. Koike, K. Yamamura, K. Inoue, T. Miyazaki, *PLoS One* **2014**, *9*, e109123.
[36] A. E. van Beek, R. B. Pouw, M. C. Brouwer, G. van Mierlo, J. Geissler, P. Ooijevaar-de Heer, M. de Boer, K. van Leeuwen, T. Rispens, D. Wouters, T. W. Kuijpers, *Front Immunol.* **2017**, *8*, 1328.
[37] W. J. Qian, J. M. Jacobs, D. G. Camp, 2nd, M. E. Monroe, R. J. Moore, M. A. Gritsenko, S. E. Calvano, S. F. Lowry, W. Xiao, L. L. Moldawer, R. W. Davis, R. G. Tompkins, R. D. Smith, *Proteomics* **2005**, *5*, 572.
[38] Z. Q. Jiang, S. B. Ying, W. Shen, X. L. He, J. Q. Chen, H. L. Xia, M. Yu, Y. Xiao, L. F. Feng, L. J. Zhu, L. Ju, X. N. Guo, Y. X. Zhang, J. W. Shen, Y. Tong, X. Zhang, J. L. Lou, *Dis. Markers* **2017**, *1725354*, 1.
[39] R. L. Gundry, Q. Fu, C. A. Jelinek, J. E. Van Eyk, R. J. Cotter, *Proteomics Clin. Appl.* **2007**, *1*, 73.
[40] G. Lucarelli, M. Rutigliano, C. Bettocchi, S. Palazzo, A. Vavallo, V. Galleggiante, S. Trabucco, D. Di Clemente, F. P. Selvaggi, M. Battaglia, P. Ditonno, *J. Urol.* **2013**, *190*, 2271.